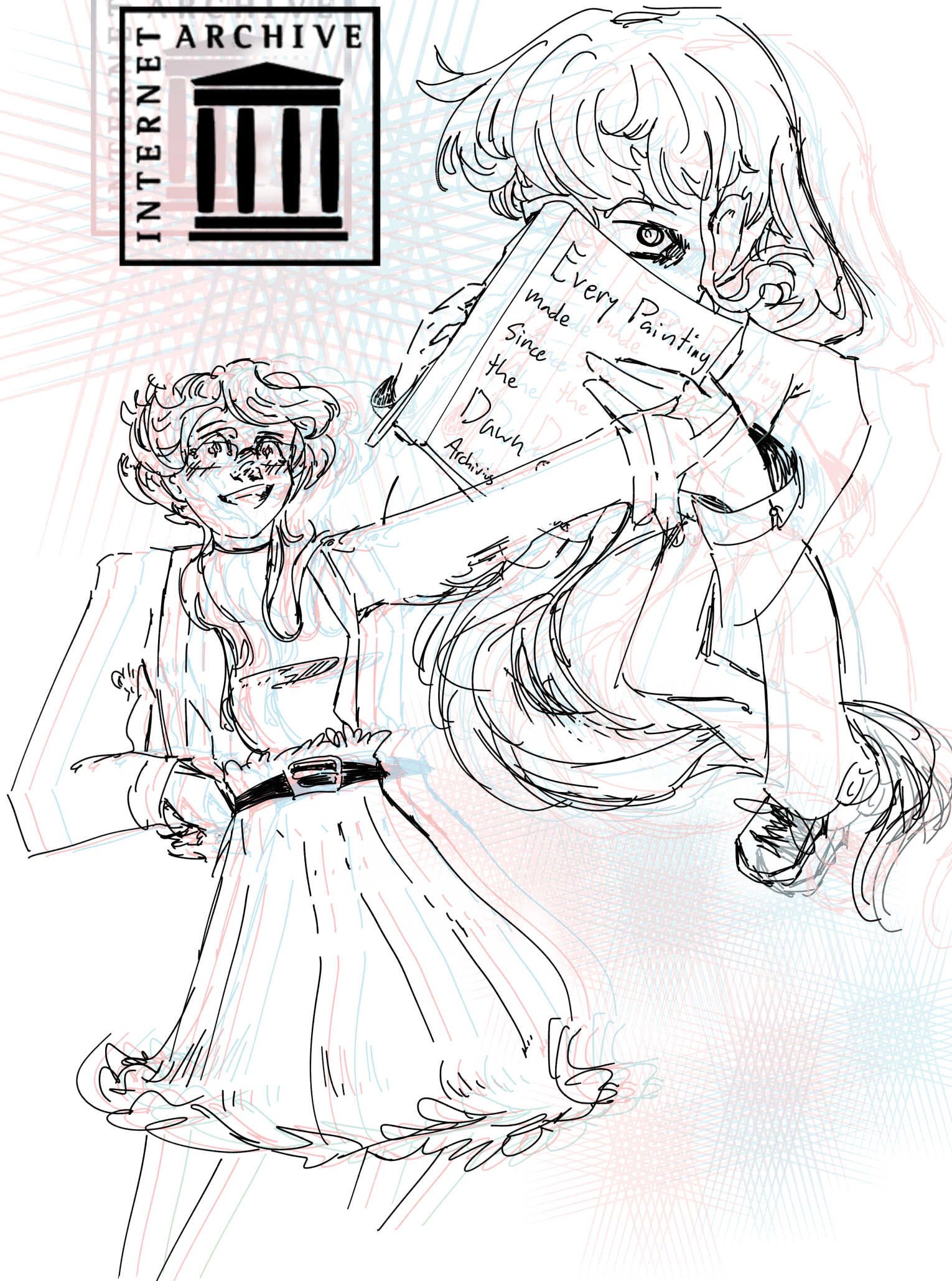


THE INTERNET ARCHIVE: THE BASICS

the only good thing about the interwebs

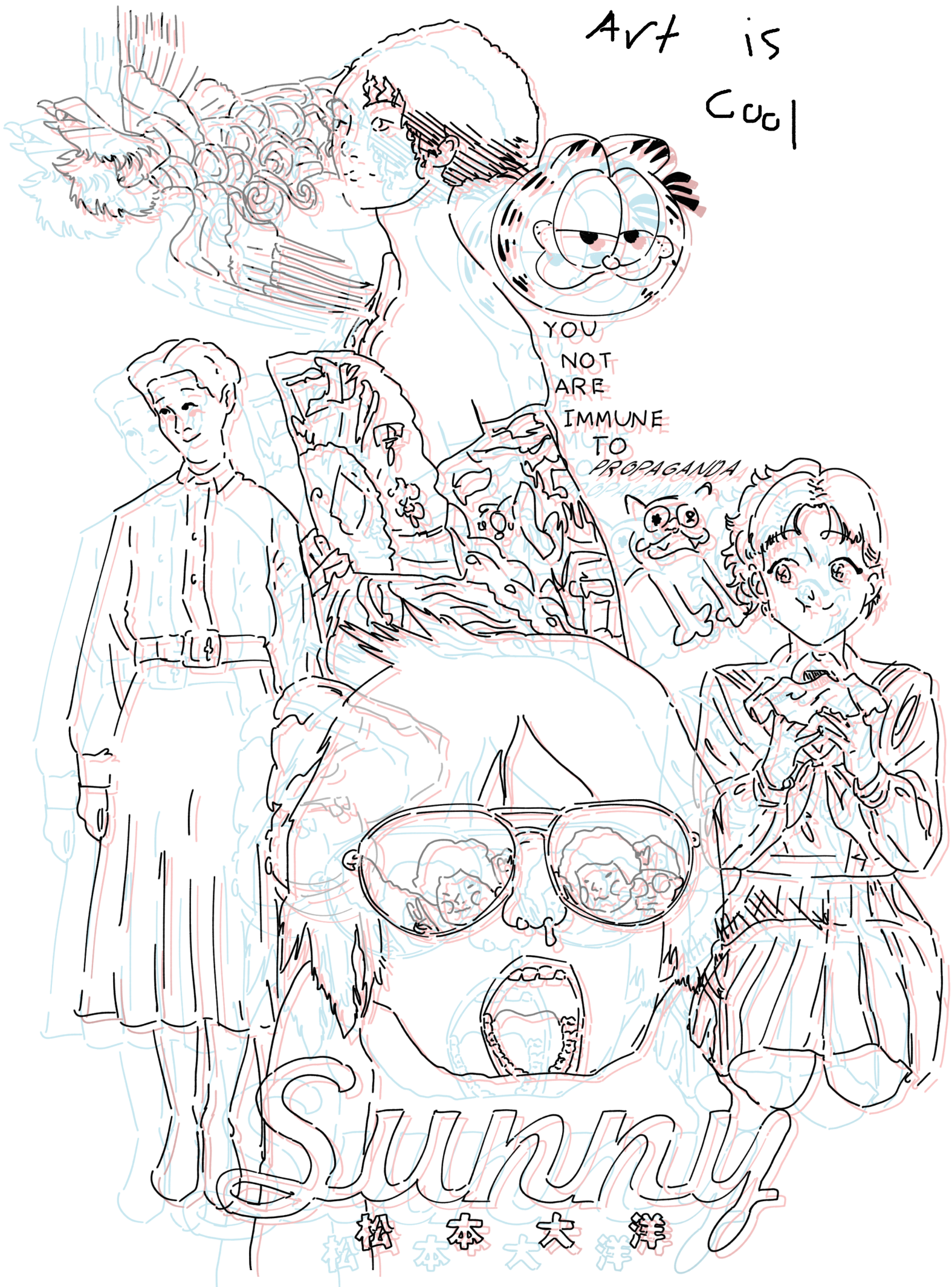


Art is
Cool

YOU
NOT
ARE
IMMUNE
TO
PROPAGANDA

Sunny

松本大洋



The Internet Archive is a deeply important tool for users of that's internet, but if you tend to stick to a few social media websites (as most people do) you probably don't understand the full importance of this archive, nor how to properly use it. But that's okay! Getting the full use out of your internet experience is a journey yet have begun to start, and the internet archive is an amazing place to!



The basics as according to archive.org/about/

- The Internet Archive is an online library meant for documenting the everyday-isms of the internet, as well as millions of older pieces of physical media.
- The Internet Archive is the home of the Wayback Machine, the primary way of seeing old versions of websites
- The Archive is home to: 916 billion web pages, 49 million books and other texts, 13 million audio recordings, 10 million videos, 5 million images, and a million software applications.
- The Internet Archive is a congressionally approved depository of government documents to make sure they're available to the public.
- If you create an account on the IA website you can upload both work you find, as well as your own work.



*The IA is also very important to areas without public libraries, without said libraries a lot of information is out of reach due to structural issues, so filling in these blind spots helps get rid of information inequality. This is also why the IA should be available in schools! Having millions of free books that's not tied to any specific district is a very valuable resource that should be available to children of all ages and zip codes. However, most schools ban the Internet archive due to it being a search engine, alongside legal obligations for schools to filter specific types of content.

How to use the Archive:

There are 5 categories of search features that the Internet Archive uses: General Metadata, Full-Text, Bookreader Text, Captions for Tv news, and the Wayback Machine which will get its own section.



Metadata is the closest to typical google search, just without a lot of the AI nonsense. Pre-Gemini boolean search also works on the archive (for those who don't know, Booleans are words like "and", "or", "to" etc.). There is also an advanced search feature that looks like this:

Advanced Search

This form allows you to perform an advanced search. You only need to fill in one field below. This can be any field. If you select "not" as your match criteria, you must select one other field.

Any field:	contains ▼	
AND Title:	contains ▼	
AND Creator:	contains ▼	
AND Description:	contains ▼	
AND Collection:	is ▼	
AND Mediatype:	is ▼	All mediatypes ▼
AND Custom field	contains ▼	
AND Custom field	contains ▼	
AND Custom field	contains ▼	
AND Date:	YYYY ▼ MM ▼ DD ▼	
AND Date range:	YYYY ▼ MM ▼ DD ▼ TO YYYY ▼ MM ▼ DD ▼	

Search

Full-text allows you to search for things that are inside of texts like books. Namely, by uploading either a PDF or an image zip file.

Bookreader is just your average text search for IA copies of books. Instead of automatically taking you to the first instance of a word/phrase, you have to click on a blue tab on the timeline in order to go to where you need to go.

Tv Caption search is a mini database of the captions of captured news stories. On the page where you find this database, you will find more specific categories to find what you need.

How to Use the Wayback Machine

The very first step to using the Wayback Machine is to find the link you want to look for. It can be copy and pasted, or just a general link like "www.somethingsomething.com". Which as I was typing this out on Google Docs, I just found out this is a real link. It's just an AI image generator for some reason.

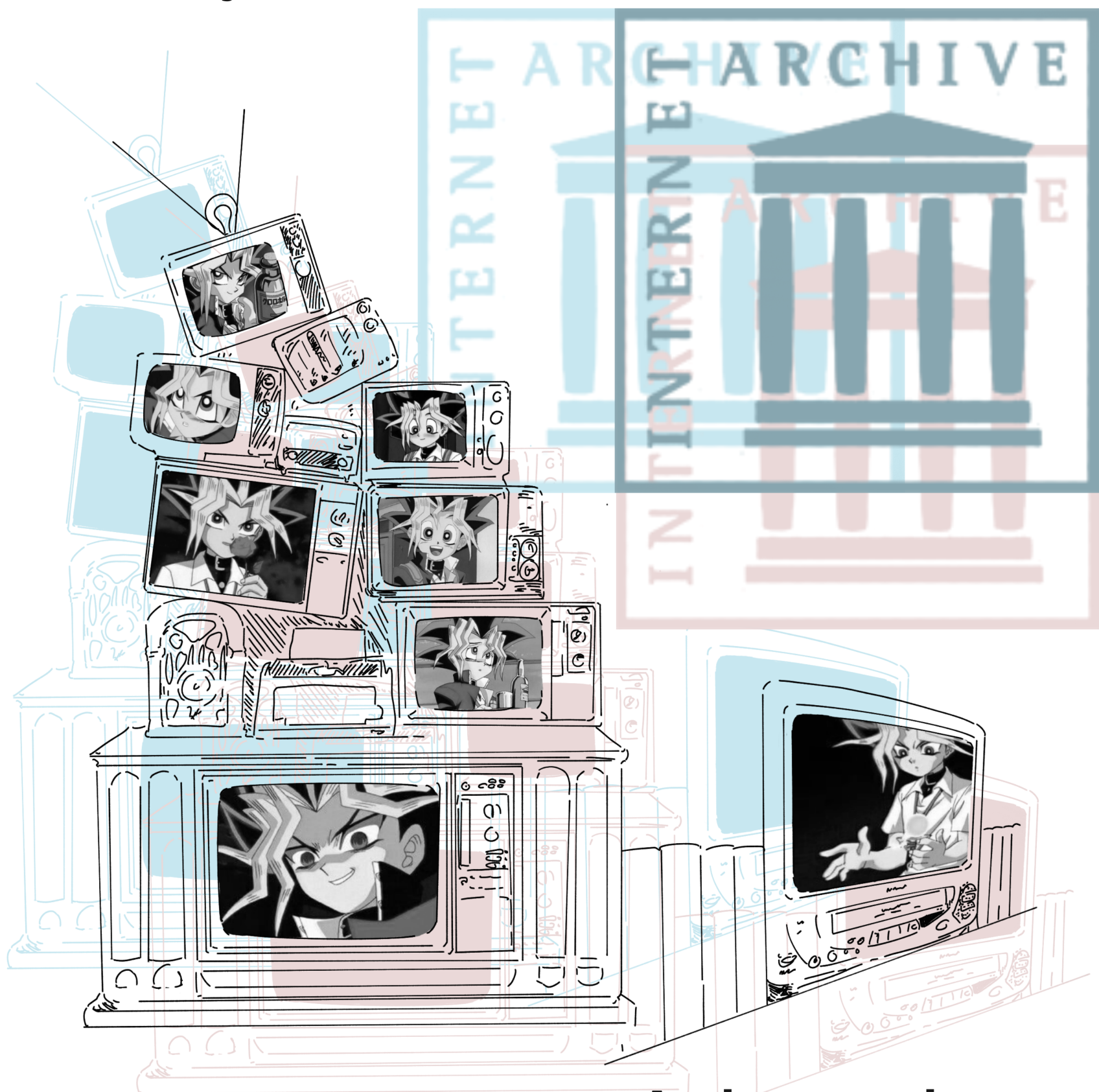
But the next step is to search by year, which is visually represented by a timeline. Click on a part of the timeline with a black bar, and the list of archived snapshots will be listed in chronological order.



If you want to contribute to the Wayback Machine, you can copy and paste a link into the "save page now" box to any website you like and preserve it! Only sites with specific security features won't work with the archives save feature. Sadly, any other links on said website won't be preserved, but you can always save those links as well.



To make this process simpler, you should download the browser extension offered on the “Save Pages in the Wayback Machine” webpage on the archive.org website. Also, if you have the time, the archive team would love for you to volunteer! There’s also a subscription service offered by the IA called Archive-It that allows you to do archive crawls without any technical knowledge if you want to make a hobby out of archiving.



And remember...

**THE INTERNET ARCHIVE IS A
NON-PROFIT AND
WOULD LOVE
YOUR
DONATIONS!!!!**



